DTIC FILE COPY

1

AD-A210 532

ARI Research Note 89-36

Battle Simulation Outcomes as Potential Measures of BCG Performance in CATTS Exercises

Gary S. Thomas
U.S. Army Research Institute

Field Unit at Fort Leavenworth, Kansas
Stanley M. Halpin, Chief

Systems Research Laboratory
Robin L. Keesee, Director

June 1989

United States Army
Research Institute for the Behavioral and Social Sciences

89 7 21 014

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON **Technical Director**	**JON W. BLADES** **COL, IN** **Commanding**

Technical review by

Lloyd M. Crumley

NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Informational Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this report are those of the author(s) and should not to be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION
Unclassified

1b. RESTRICTIVE MARKINGS
--

2a. SECURITY CLASSIFICATION AUTHORITY
--

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE
--

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release;
distribution is unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
ARI Research Note 89-36

5. MONITORING ORGANIZATION REPORT NUMBER(S)
--

6a. NAME OF PERFORMING ORGANIZATION
U.S. Army Research Institute for the Behavioral and Social Sciences

6b. OFFICE SYMBOL *(If applicable)*
PERI-SL

7a. NAME OF MONITORING ORGANIZATION
--

6c. ADDRESS *(City, State, and ZIP Code)*
ARI Field Unit at Fort Leavenworth, KS
P.O. Box 3407
Fort Leavenworth, KS 66027-0347

7b. ADDRESS *(City, State, and ZIP Code)*
--

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
U.S. Army Research Institute for the Behavioral and Social Sciences

8b. OFFICE SYMBOL *(If applicable)*
PERI-S

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
--

8c. ADDRESS *(City, State, and ZIP Code)*
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
62785A	791	1301	H2

11. TITLE *(Include Security Classification)*
Battle Simulation Outcomes as Potential Measures of BCG Performance in CATTS Exercises

12. PERSONAL AUTHOR(S)
Thomas, Gary S. (ARI) (Combined Arms Tactical Training Simulator)

13a. TYPE OF REPORT
Final

13b. TIME COVERED
FROM 1982 TO 1983

14. DATE OF REPORT *(Year, Month, Day)*
1989, June

15. PAGE COUNT
24

16. SUPPLEMENTARY NOTATION This report was originally published in 1983 as an internal ARI document not available for general distribution. It has been made available because it contains data now of general interest to C2 researchers.

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Command; Command and control; Control; Staff evaluation; Battle outcome; Performance

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)* Ten Battalion Command Groups consisting of a battalion commander, an S1, S2, S3, S4, four company commanders, and roughly ten supporting staff members were exercised on the CATTS battle simulator. CATTS provides a considerable quantity of battle status data so it is possible to compute loss exchange ratios, relative exchange ratios, surviving maneuver force ratio differentials, command and control index of lethality ratios, and change in combat ratios--all of which are considered battle outcome measures when computed at the end of the battle. Results showed that, although all measures were calculated from the same battle data, correlations between them were not always high (.96 to .02) and were sometimes negative (-.35 to -.19) Attempts to relate performance measures to battle outcome measures were unsuccessful and in the case of subjective evaluations by controllers, there was reason to believe that the data were contaminated by the controllers' dual roles as evaluator and training event controller. More research and higher quality data will be required to better evaluate the relationship between battle outcome measures, and between those measures and task performance.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL
Lloyd M. Crumley

22b. TELEPHONE *(Include Area Code)*
(913) 684-4933

22c. OFFICE SYMBOL
PERI-SL

DD Form 1473, JUN 86 *Previous editions are obsolete.*

BATTLE SIMULATION OUTCOMES AS POTENTIAL MEASURES OF BCG PERFORMANCE IN CATTS EXERCISES

CONTENTS

LIST OF TABLES

CONTENTS (Continued)

Page

LIST OF TABLES (Continued)

LIST OF FIGURES

Battle Simulation Outcomes as Potential Measures of BCG Performance in CATTS Exercises

In the past several years battle simulations have gained increased credibility as systems for training command and control (C^2) processes at tactical echelons from battalion through corps. Of the battle simulations developed to support C^2 training, the most sophisticated in terms of the extent of automation and, therefore, ability to represent battlefield events in real time, is the Combined Arms Tactical Training Simulator (CATTS). CATTS is used to train battalion command groups (BCG) and serves as a test bed to improve training procedures and to specify requirements for future simulations. In order to address training issues within the context of the CATTS system, it is necessary first to develop measures of BCG performance, so that changes in C^2 performance can be determined and used to assess the training value of various CATTS characteristics and procedures.

During the last several years, ARI has conducted research directed toward developing a variety of BCG performance measures in CATTS exercises. These efforts have resulted in a questionnaire based on a modified form of the Army Training and Evaluation Program (ARTEP) (Barber and Kaplan, 1979; Kaplan and Barber, 1979) that has been used to assess the extent to which ARTEP subtasks are exercised in CATTS. The BCG players participating in CATTS exercises and the CATTS control staff have rated player performance on ARTEP tasks and subtasks using this questionnaire. Also developed by ARI and the Automated Command Training Division (ACTD) is an information flow questionnaire designed to determine which players received and transmitted information required by the group to effectively plan and execute a mission during CATTS exercises (Kaplan, 1980). These instruments were designed primarily to assess the process by which BCGs coordinate and implement a battle plan, and do not directly assess the consequences of this process. The following discussion considers the usefulness of simulated battle outcomes as potential measures of the consequences of these BCG processes.

Simulation outcomes could complement the performance measures already being collected and could be especially valuable due to their less subjective nature. There are, however, several questions associated with using battle outcomes as performance measures in CATTS that have prevented their utilization in the past. First, to what degree are the outcomes of simulated battle engagements realistic; and second, to what degree are the outcomes a valid indication of BCG performance? The validity question points to the need for some criterion measure of BCG performance that simulation outcomes can be validated against, or for a procedure for attributing validity to the battle outcomes.

The deterministic nature of the simulation outcomes can be addressed in terms of the CATTS software and internal algorithms. Such variables as weapons effects, priority of fire, movement rates, and loss probabilities are all programmed into the CATTS system. The computer takes from CATTS controllers, input based on C^2 decisions by the BCG, and determines the battlefield outcomes as a result of those inputs. The battle simulation outcomes are largely deterministic rather than probablistic, since replications of the same maneuvers and engagements result in very similar friendly and OPFOR losses.

TRASANA (1979) has investigated the battle calculus, movement routines, line of sight calculations, weapons effects curves, etc., of CATTS and has determined them to be realistic representations of actual battle. In addition, continuous efforts have been made by ACTD over the last several years to further refine the CATTS algorithms, so that more realistic outputs are derived from the inputs to the system.

Currently there is no absolute criterion measure of BCG performance available to which the simulation outcomes from CATTS can be compared. But, in terms of validity, the outcomes of battle should: (1) reflect expected differences in battle outcomes for different types of missions (e.g., covering force vs. attack); (2) reflect expected outcomes due to initial combat ratio (relative combat power) of OPFOR vs. friendly forces; and (3) reflect a systematic relationship to subjective ratings of BCG performance by CATTS controllers. The current paper is an initial examination of the validity attributable to simulation outcomes as measures of BCG performance.

METHOD

Participants

Participants were ten BCGs consisting of a battalion commander, an S1, S2, S3, S4, four company commanders, and roughly ten supporting staff members. The BCGs exercised in CATTS between October 1980 and October 1981. BCGs were either mechanized infantry (n=4), armored cavalry (n=3), or light infantry (n=3), that conducted, on either Fulda or Sinai terrain, a covering force (mech and cav) operation or a static defense (infantry only), followed by an attack on the next day. All mechanized and cavalry units and one infantry unit exercised on the Fulda Gap terrain. However, the terrain was slightly different for each unit type. The remaining two infantry units exercised on a Sinai terrain. The OPFOR typically deployed in basically the same fashion in all equivalent missions (attack or defense). BCGs conducted missions in conjunction with adjacent battalions controlled by a brigade staff, played by CATTS controllers.

CATTS Exercises

On the first day of CATTS exercises, BCG staffs received the brigade operations order from CATTS controllers playing the roles of brigade S1, S2, S3, and S4. While the BCG staffs were planning their operations orders, CATTS controllers, functioning as maneuver controllers, fire support officers, and computer interactors trained BCG company commanders and their fire support personnel to participate in CATTS exercises. On day two, BCG players conducted a defensive mission (covering force or static defense). At the conclusion of the exercise, CATTS controllers rated the performance of the BCGs. On the third day, BCGs received a frag order, conducted an attack, and were again rated by CATTS controllers.

Battle Outcome Components

Computer-generated summary reports of battle status of opposing forces provided primary data for this investigation. These reports indicated initial level of equipment in all OPFOR and friendly units, plus a summary of those pieces of equipment lost by opposing units through the course of the battles. CATTS battle calculus included a relative weight for each weapon system according to its degree of fire power. For example, an M60A1 tank had a weight of 73, whereas an M113 APC had a weight of 19. The weighting system contributed to the computer decision rules that determined which types of equipment fired at and destroyed other equipment types. This weighting system was used in the current research to determine relative initial strength and the relative battle losses of OPFOR and friendly forces. This was accomplished by multiplying the weighting factors by the number of the corresponding equipment types under consideration. The products were then summed across all OPFOR units and friendly forces, respectively, that were present on the battlefield. The total weighted, initial strength of OPFOR and friendly forces appear in columns one and two of Table 1. The table also indicates the initial combat ratio (OPFOR initial forces divided by initial friendly forces) and the total losses sustained by opposing forces in the simulated battles. Appendix A includes the rationale for determining initial force levels.

Subjective Ratings

The other source of data was CATTS controller ratings of overall BCG performance on each exercise day. The number of controllers who rated BCG performance in each CATTS exercise ranged from five to nine depending upon the availability of experienced controllers. Since the rating instrument used a magnitude estimation technique (Stevens, 1973), raw ratings were converted to $\log_{10}$ to normalize the distribution of scores. Those logs were further converted to Z-scores for each individual rater in order to control for systematic rater-response bias. A mean standardized performance score was then computed for each BCG on each exercise day.

TABLE 1

Weighted Levels of Forces and Losses

Unit	Initial OPFOR	Initial Friendly	Initial Combat Ratio	OPFOR Loss	Friendly Loss
Covering Force/Defense					
Mech 4	18326	6130	2.990	3883	2606
Mech 5	18326	6345	2.880	3965	3046
Mech 6	18326	6191	2.960	2682	2654
Mech 10	18326	6191	2.960	2745	2802
Cav 2	18326	9663	1.897	5981	5063
Cav 3	18326	9082	2.018	5101	5040
Cav 4	18326	9030	2.029	6066	4634
Inf 1	18326	11804	1.553	5449	5495
Inf 8	18326	15935	1.150	5760	8054
Inf 9	18326	8558	2.141	4572	4274
Attack					
Mech 4	11916	5495	2.169	1993	1511
Mech 5	11916	5768	2.066	428	1825
Mech 6	11916	5664	2.104	682	1094
Mech 10	11916	5664	2.104	1809	2434
Cav 2	13511	12305	1.098	1547	4547
Cav 3	13511	12795	1,056	2064	4198
Cav 7	13511	12761	1.059	824	2141
Inf 1	10729	11800	.909	4338	5631
Inf 8	10729	14488	.741	1491	3300
Inf 9	17304	6096	2.839	814	1147

Procedure

The battle outcome components compiled in Table 1 for all BCGs participating in defensive and attack missions served as data to compute the simulation outcomes as potential measures of battlefield performance. These include loss exchange ratio (LER), a relative exchange ratio (RER), and surviving maneuver force ratio differential (SMFRD), which have all been used in previous research and/or combat modeling (USACDC, 1973). All these measures are mathematical relationships of OPFOR and friendly losses or surviving forces at the end of an entire battle. The mathematical formulas used to compute these measures appear in Appendix B. The LER is simply computed by dividing the weighted total OPFOR loss by friendly weighted total losses; RER is a ratio of the percentage of OPFOR losses to the percentage of friendly losses; and SMFRD is the percent of friendly forces surviving minus the percent of OPFOR surviving battle. Two other potential measures of battle performance were generated to predict overall controller ratings of BCG performance: the Command and Control Index of Lethality Levels (C^2ILL) ratio and the change in combat ratio (ΔCR). Calculations of these also appear in Appendix B.

All other things being equal, it is preferable to have a higher proportion of friendly forces surviving battle and to destroy a high proportion of enemy weapon systems. This assumption lead to the formulation of the C^2ILL ratio, which was calculated by adding the percentage of OPFOR losses to one-half the percentage of friendly forces surviving. Preliminary data analysis indicated that in assessing player performance, the controllers apparently were placing more emphasis on OPFOR attrition than on the proportion of friendly forces lost or surviving. Therefore, in the current formula, percent of friendly forces was divided in half, so that it deemphasized friendly surviving with respect to OPFOR losses. It may be noted that the formulas for C^2ILL and SMFRD are similar with the exception of the above weighting factor applied to friendly forces surviving.

The other potential measure of battlefield performance (ΔCR), was formulated on the assumption that it is preferable to end the battle with a more advantageous combat ratio than existing prior to battle. As indicated in Table 1, OPFOR typically had a combat power advantage prior to battle as depicted by initial combat ratios in excess of 1.0. The degree to which the friendly forces can change that combat ratio during battle to a more favorable ratio, should be an indication of how well they employed weapon systems against the OPFOR. This change in combat ratio (ΔCR) was calculated by subtracting the end of battle combat ratio from the initial combat ratio and dividing the difference by the initial combat ratio. Higher positive ΔCR's indicate better battlefield performance. Higher positive values are also preferable on the other measures of battlefield performance. The purpose of formulating these last post-hoc measures was to find additional measures that correlated with controller ratings, and that generalized across more mission and unit types.

RESULTS

The present research, which was conducted in a very uncontrolled environment, attempts to formulate battle simulation outcomes that bear some relationship to controller ratings of C^2 performance. Several preliminary analyses are presented to arrive at a better understanding of the components of these simulation outcomes, particularly losses and initial strengths of opposing forces. These analyses are of interest since the values of these components determine the magnitude of the simulation outcomes; and consequently, the relationship between outcomes and controller ratings. The following discussion also identifies some uncontrolled variables (e.g., initial combat ratio) that could affect simulation outcomes and the relationship between outcomes and ratings.

As stated previously, battle outcomes (e.g., losses) should logically be affected by such variables as mission type and initial combat ratios. But, for purposes of performance evaluation and training, it is desirable to develop measures of battlefield performance that generalize across such variables. The first analysis investigates the impact of some of these variables on controller ratings of performance, so that results can then be compared to similar analyses performed on battle outcomes.

Controller Ratings Of Performance

The controller performance ratings (Appendix C) were analyzed in a 2X3 ANOVA to determine if the ratings varied as a function of two types of mission or three types of unit. The mean values of performance ratings for mission and unit type are presented in Figure 1. Although ratings were higher for Cav and Inf than for Mech, the results of the ANOVA only approached significance ($F_{2,7} = 3.953$, $P < .10$). There was no difference in ratings for mission, however, the unit by mission interaction was significant ($F_{1,7} = 7.169$, $P < .05$). This interaction apparently reflected the fact that performance ratings increased from defense to attack for Mech units only, while both Inf and Cav units' performance ratings declined somewhat from defensive to attack missions. Of particular interest is the fact that Mech units were rated exceptionally low in defensive missions. This will be the topic of later discussion.

OPFOR And Friendly Losses

Since controller ratings of BCG performance were influenced by the mission and unit type variables, it is possible that OPFOR and friendly losses in CATTS exercises were also affected by these variables. The potential effects are of interest, since equipment losses are the major components of the simulation outcome measures to be discussed later.

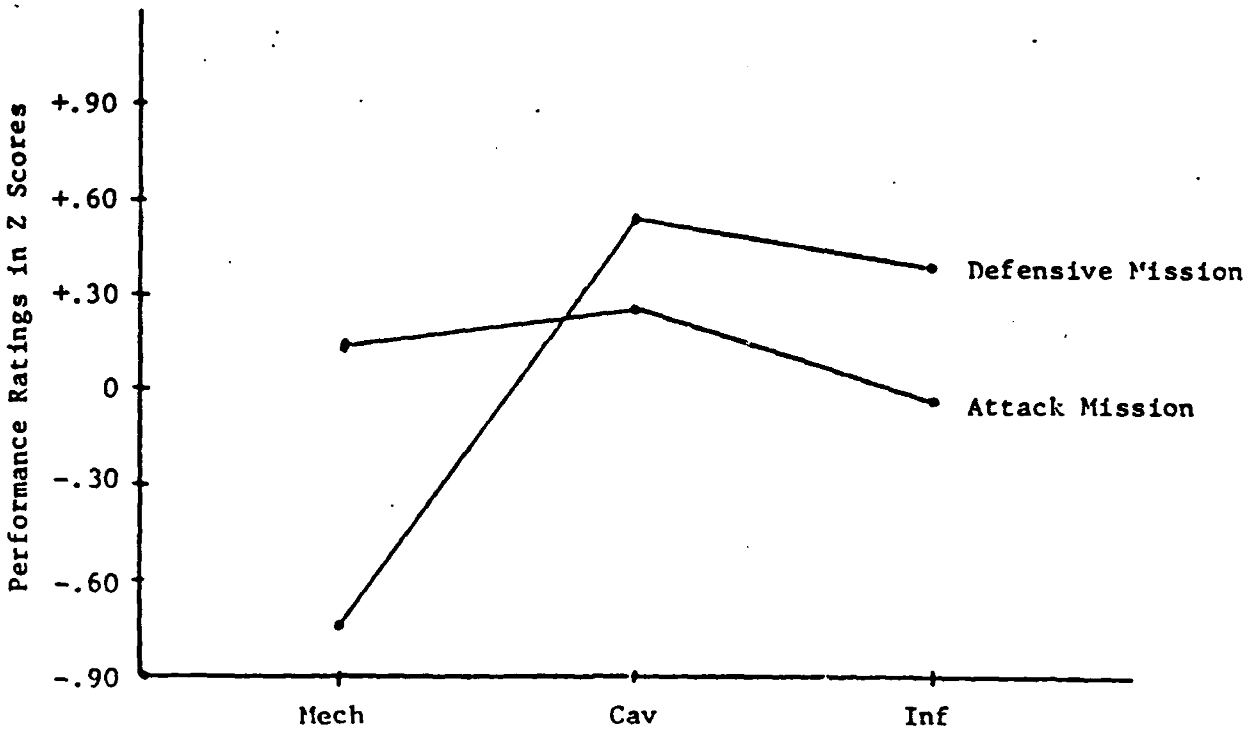

Figure 1. Mean controller performance ratings for Mech, Cav, and Inf units conducting defensive and attack missions.

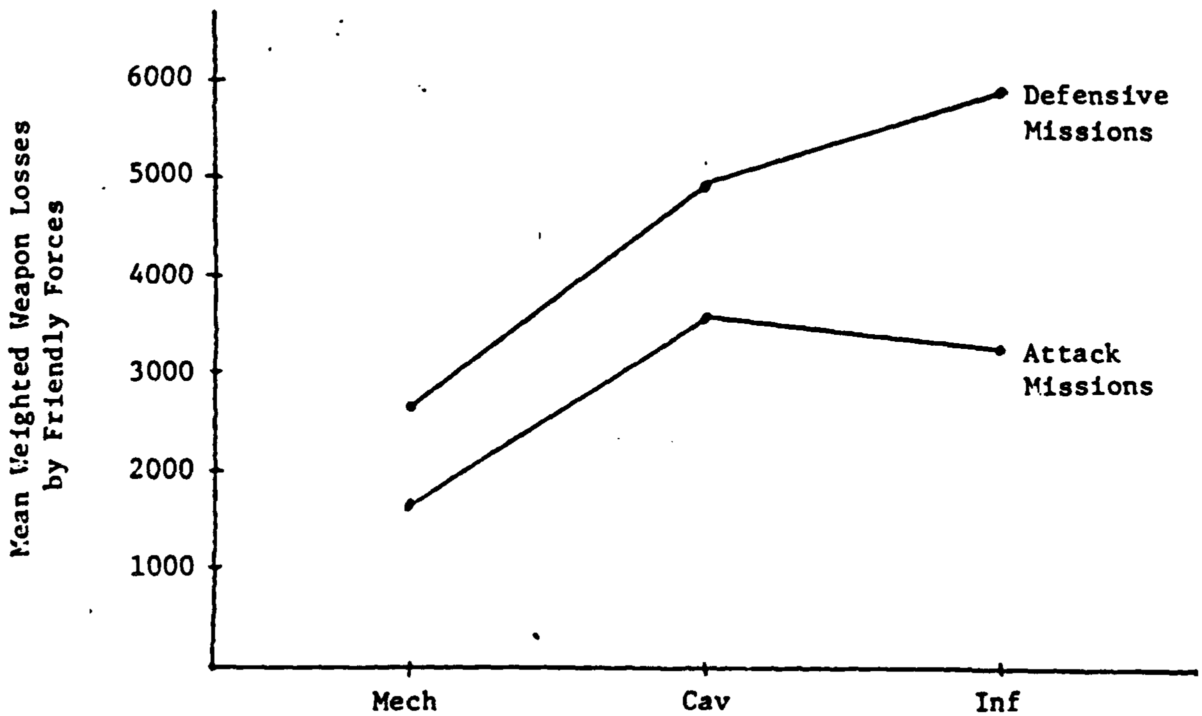

Figure 2. Mean losses received by friendly units for defensive and attack missions.

All else being equal, it can be argued that units which attrit large numbers of enemy equipment while minimizing losses, should typically be perceived as performing better than battalions that do not. The following set of analyses were conducted to determine the influence of mission and unit type on friendly and OPFOR losses. The mean values for total losses incurred by friendly units in the two mission types are presented in Figure 2. An analysis (2X3 ANOVA) of friendly losses as a function of mission and unit type revealed significant main effects for both variables, ($F_{1,7}$= 11.356, $P < .025$ and $F_{2,7}$ = 5.836, $P < .05$), respectively. Fewer friendly losses were sustained in attack missions and by Mech units. The mission effects were reasonable, since friendly units typically encounter far less OPFOR in attack scenarios. The fact that Mech units sustained less casualties and received lower ratings than Inf and Cav units appears inconsistent with the above hypothesis. To be consistent with the hypothesis and the lower performance ratings, Mech units would have had to inflict considerably fewer losses on the OPFOR to compensate for the low level of losses they sustained.

Analysis of OPFOR losses (2X3 ANOVA) does not resolve this inconsistency, since no significant differences for unit type were observed. The main effect for mission type was, however, significant ($F_{1,7}$ = 54.44, $P< .005$), where more OPFOR equipment was destroyed in the defensive missions than in attacks. In an attempt to explain the lack of expected differences in OPFOR losses inflicted by Mech units in comparison with Cav and Inf units, it was observed that occasionally during Mech covering force missions, excessive air missions had to be initiated by friendly brigade controllers to prevent friendly battalions from being overrun by OPFOR. Since controllers appeared to be sensitive to these situations, and adjusted their ratings accordingly, the friendly Mech data were "adjusted" in subsequent analyses. Units whose defensive air "kills" (Mech 5 and Mech 10) exceeded those attained by all other units by one standard deviation from the mean for all units (Mech, Cav, and Inf), had that amount of "kills" subtracted from scores of OPFOR losses induced. These "adjusted" values for OPFOR losses inflicted by friendly units are presented in Figure 3.

A subsequent reanalysis (2X3 ANOVA) of OPFOR losses reveal a significant main effect for unit type ($F_{2,7}$ = 4.985, $P< .05$), where Cav and Inf units achieved more OPFOR "kills" than did Mech units. The "statistical adjustment" described above results in findings that are consistent with controller ratings of performance, since one of the main objectives of defensive operations is to attrit the enemy. It is also possible that poor deployment of existing assets contributed to both low OPFOR attrition and low friendly casualties sustained by Mech units, especially in the covering force mission.

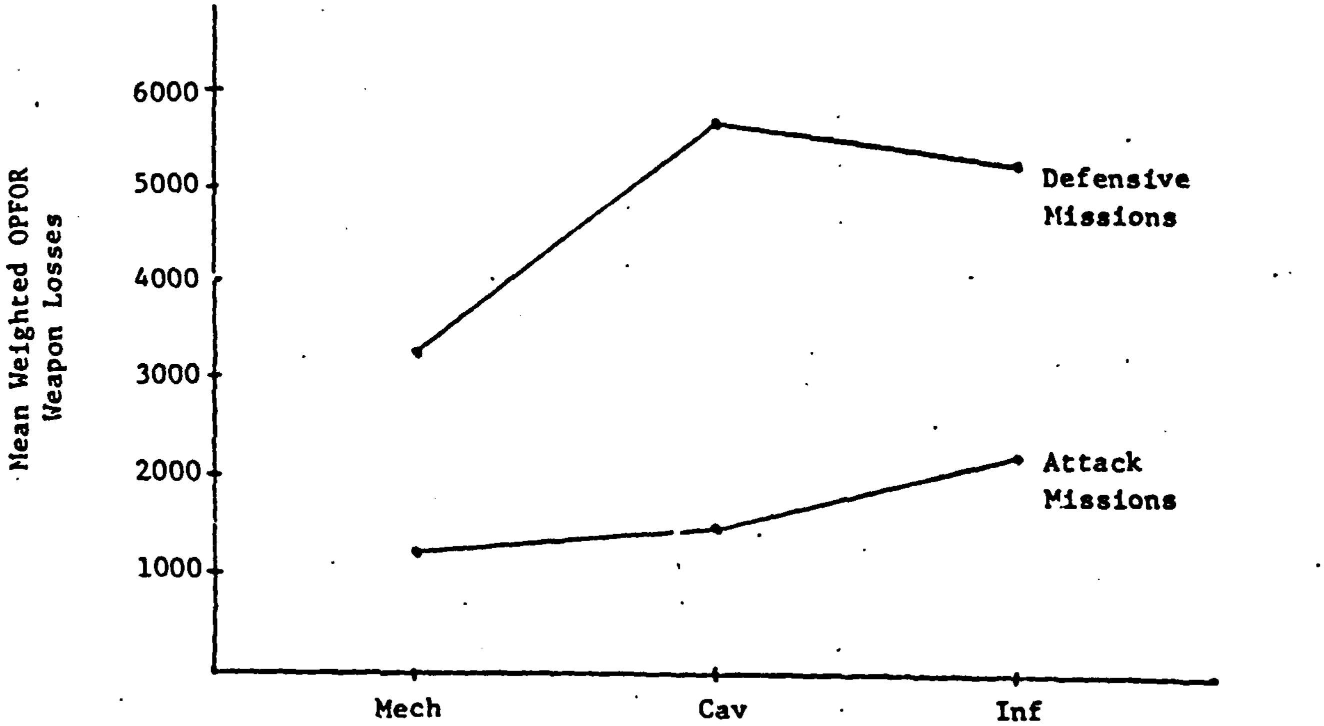

Figure 3. Mean OPFOR losses inflicted by friendly units for defensive and attack missions.

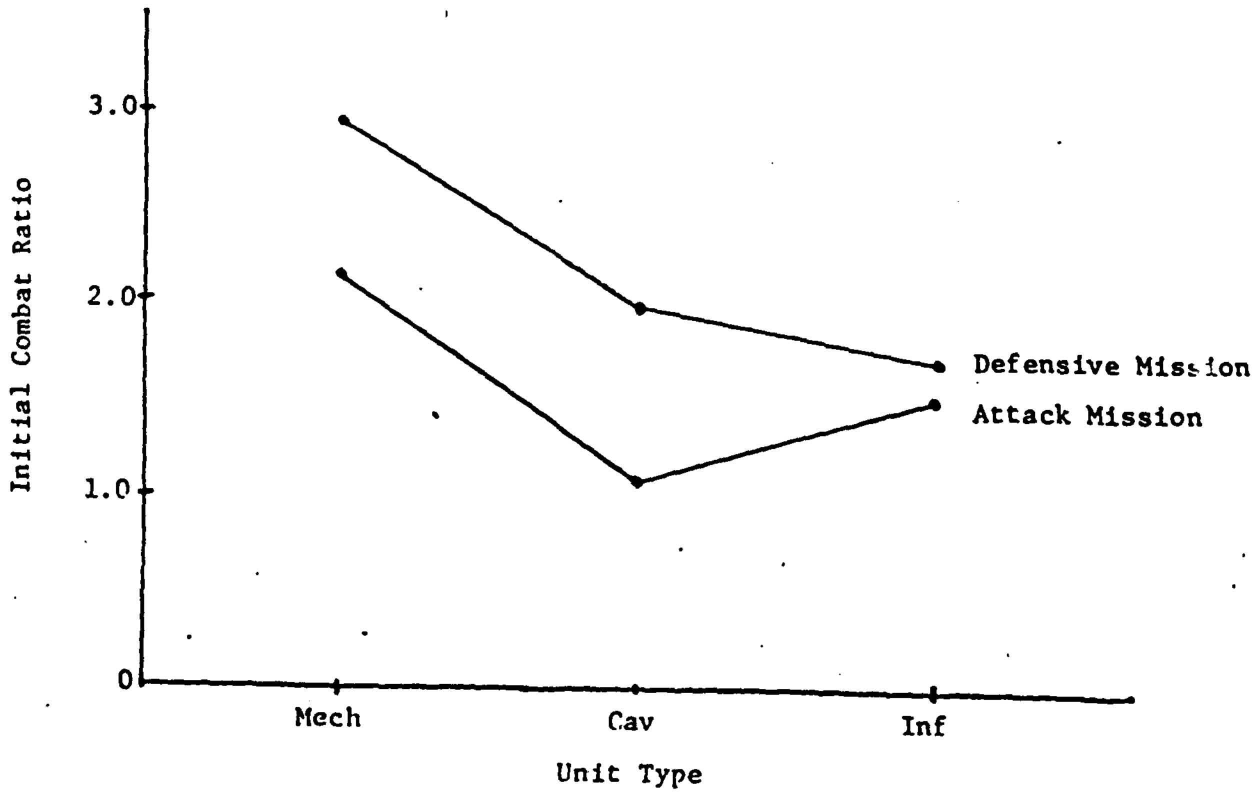

Figure 4. Mean initial combat ratio for Mech, Cav, and Inf units conducting defensive and attack missions.

Initial Combat Ratio

Since the factors unit and mission type apparently affect both performance ratings and the outcomes of battle, it is desirable to identify underlying variables which may impact on both of these factors. One such variable is the combat ratio that exists prior to battle. It is reasonable to assume that units with more initial strength relative to the OPFOR are likely to perform better on the battlefield; therefore, analyses were conducted on initial combat ratio and its relationship to unit and mission type. The mean values tested for significance represent the proportion of OPFOR to friendly forces at the beginning of battle and are presented in Figure 4. Results of the analyses (2X3 ANOVA) indicated a reliable difference for unit type ($F_{2,7} = 6.047$, $P < .05$), where combat ratio favoring OPFOR tended to be higher for Mech than for Cav and Inf. Not surprisingly, mission type was also a significant variable ($F_{1,7} = 27.627$, $P < .005$), where combat ratios favoring OPFOR were higher in defensive missions.[1]

Mech units tended to be rated lower by controllers and also to be exposed to higher and less favorable combat ratios. In fact, a correlation of $-.67$ ($P<.001$), indicated that as combat ratio increased (became less favorable to friendly forces), controller ratings of performance tended to decrease regardless of mission or unit type. Battle outcomes (losses) were also likely affected by prebattle combat ratio; for example, Mech units sustained and inflicted fewer losses and were exposed to poorer initial combat ratios. Eventhough losses may be related to controller ratings of performance, it is apparent that other factors contributed to these ratings. For example, battle losses and initial combat ratios both varied as a function of mission type, but mission type did not significantly influence controller ratings. It is, therefore, possible that controllers adjusted their performance ratings to take into account mission differences, such as the fact that the defenders had a tactical advantage and could better fight outnumbered. In addition, the way in which BCGs deployed their assets should in great degree determine how well they performed on the simulated battlefield.

As predicted, it is apparent that the outcomes of (e.g., losses) are affected by such variables as unit and mission type, which are confounded with initial combat ratios. There is also some degree of relationship between these variables and controllers ratings of performance. Since it is desirable to assess simulated battle performance across these variables, the following discussion emphasizes potential ways of combining battlefield losses into composite scores (simulation outcomes) in order to predict C^2 performance ratings.

[1] As indicated in Appendix A, combat ratio as a function of mission type does not totally reflect doctrine since the figures only reflect combat power available on the battlefield and not the way in which it was employed.

Simulation Outcomes And Performance Ratings

Investigations into the relationship between potential measures of BCG performance were conducted as a function of friendly unit and mission type. The potential measures of simulated battle performance were correlated with controller ratings of performance (Appendix C) to arrive at a direct comparison of the relationships among these variables, as shown in Tables 2 through 4.

Table 2 is a matrix that indicates the situations in which the rank order of controller ratings of BCG performance was identical to the rank order of simulation outcomes for each unit type within each mission type. Rank order of these variables were used since n-size in each situation was too small to calculate any meaningful correlations. As indicated in the table, controller ratings are predicted by the rank order of simulation outcomes in 25 of 30 possible comparisons. Only the C^2ILL ratio rankings perfectly predicted the rank of controller ratings in all six unit by mission situations.

Table 3 represents a correlation matrix where the five candidate measures of battle performance were compared to controller ratings of performance for each type of unit (Mech, Inf, and Cav), calculated across mission type. As indicated by the table, each candidate measure correlates significantly with performance ratings for one or more of the unit types, but not for all three. It appears that the degree of relationship is unit dependent.

Table 4 indicates those battle performance measures that correlate significantly with controller ratings across unit type when mission (defensive vs. attack) is considered separately. Significant correlations appear in the table for all comparisons except ΔCR in defensive missions and LER in both defensive and attack missions. RER and C^2ILL appear to be particularly good predictors when mission is held constant.

The other correlations appearing in Table 4 are between controller ratings of performance and the five simulation outcome measures for all 19 exercises regardless of mission or unit type. As indicated by the table, RER, SMFRD, and C^2ILL correlate significantly with controller ratings. The highest correlations obtained were for the C^2ILL and SMFRD. These separately account for nearly 50% of the variance in controller ratings, and combined in a multivariate analysis, account for 66% of the variance. Apparently the factors including unit and mission type still impact on these measures, since the magnitude of these correlations is not as high as those derived when mission and unit type are controlled. There is also no simulation outcome that appears to be the "best" predictor of controller ratings in all situations, although C^2ILL appears to be the most consistant.

Table 2

Rank Order Comparisons of Simulation Outcomes and Controller Ratings of Performance for Units and Missions

	LER	RER	SMFRD	ΔCR	C^2ILL
Mech covering force (n=4)	X	X	X	X	X
Cav covering force (n=3)	X	X	X	X	X
Inf defense (n=3)		X	X	X	X
Mech attack (n=4)	X		X	X	X
Cav attack (n=3)	X	X			X
Inf attack (n=2)*		X	X	X	X

X -Identifies situations where the rank order of simulation outcomes were identical to the rank order of controller ratings of BCG performance.

* -Indicates that one infantry unit in the attack was excluded from this and all subsequent analyses due to the fact that the unit failed in its mission. Only this unit totally failed in its mission resulting in extraordinary low controller ratings, but relatively high simulation outcome scores. Controllers were apparently responding to the fact that mistakes made by the unit required adjacent units to complete the mission.

Table 3

Correlations Between Simulation Outcomes and Performance Ratings for Unit Type Combining Across Mission Type

	LER	RER	SMFRD	ΔCR	C^2ILL
Mech (n=8)	r=-.027	r=.317	r=.831[b]	r=.956[c]	r=.511
Cav (n=6)	r= .812[a]	r=.918[c]	r=.139	r=-.059	r=.949[c]
Inf (n=5)	r=-.420	r=.874[a]	r=.429	r=.476	r=.820[a]

[a] Pearson correlation with $P < .05$
[b] Pearson correlation with $P < .005$
[c] Pearson correlation with $P < .001$

Table 4

Correlations Between Simulation Outcomes and Performance Ratings for Mission Type Combining Across Unit Type and for All Exercises Combined

	LER	RER	SMFRD	ΔCR	C^2ILL
Defensive missions (n=10)	r=-.009	r=.946[c]	r=.800[b]	r=.200	r=942[c]
Attack missions (n=9)	r= .466	r=.901[c]	r=.827[b]	r=.764[a]	r=.953[c]
All missions and unit types (n=19)	r=-.079	r=.590[b]	r=.696[c]	r=.386	r=.697[c]

[a] Pearson correlation with P <.01
[b] Pearson correlation with P <.005
[c] Pearson correlation with P <.001

Table 5

Intercorrelations Between Candidate Measures of Battlefield Performance

	LER	RER	SMFRD	ΔCR	C^2ILL
LER	-	r=.547[b]	r=-.192	r=-.347	r=.475[a]
RER		-	r= .372	r= .021	r=.962[c]
SMFRD			-	r= .764[c]	r=.471[a]
ΔCR				-	r=.094
C^2ILL					-

[a] Pearson correlation with P <.05
[b] Pearson correlation with P <.005
[c] Pearson correlation with P <.001

Finally, intercorrelations between the five candidate measures of battlefield performance, regardless of mission or unit type, are presented in Table 5. As indicated by the table, there are several instances where the measures are related, but only RER and C^2ILL are nearly perfectly correlated. It may be, therefore, that these measures are in general assessing somewhat different aspects of battlefield outcomes.

GENERAL DISCUSSION

This initial attempt at identifying simulated battle outcomes as measures of BCG performance appeared to be successful. The procedure of assigning fire power weights, which are part of the CATTS battle calculus, to individual weapons systems allowed calculations of total pre-battle combat power and of total losses through the course of simulated battle. The battle losses were responsive to both mission and unit type variables. Where few units are actually engaged, attack missions resulted in significantly fewer losses for both OPFOR and friendly units. Mech units also inflicted fewer casualties on OPFOR than did Cav and Inf units. This may have been due to the fact that Mech units had less fire power and faced significantly less favorable initial combat ratios. These results are consistent with expectations in real combat. The inability of Mech units to inflict high casualties on the OPFOR in the covering force mission appears to be reflected in exceptionally low controller ratings for this condition. It should be pointed out that across all defensive missions, the correlation between performance ratings and absolute number of OPFOR losses were $r=.96$. However, a comparable correlation for attack missions was only $r=.08$, indicating that different factors relate to controller ratings depending upon mission type.

The simulation outcomes, which are composites of friendly and OPFOR losses, were typically good predictors of controller ratings within some mission and unit type constraints. The LER only correlated with ratings in some cases where mission and/or unit type was held constant. On the other hand, RER, SMFRD, and C^2ILL correlated very highly with ratings across unit type for both attack and defensive missions. And, when the data was combined across all conditions, SMFRD and C^2ILL both predicted overall performance ratings to a high degree ($r=.70$).

These results are highly promising with respect to the potential use of simulation outcomes as measures of BCG performance on the simulated battlefield in CATTS. However, in the current investigation the simulation outcomes were calculated to maximize the relationship with performance ratings. Therefore, a shrinkage of the observed relationships could be expected in replications of the research. On the other hand, the correlations could improve if factors such as initial combat ratio were controlled or systematically manipulated.

Another factor impacting the battle outcomes is the influence of CATTS controllers performing the roles of brigade staff members. Situations where their influence could be observed were in brigade initiated airstrikes, in artillery support, and in assistance given to battalion by adjacent units. It is desirable that these factors be controlled in future research to separate the influence of CATTS controllers playing brigade staff from the performance of BCG on the battlefield.

Variables such as terrain and factors other than losses which contribute to overall mission accomplishment could not be systematically studied in the present descriptive investigation. No highly significant correlations between ratings and simulation outcomes were observed for Inf units. This may have been due in part to the fact that Inf units exercised on both Sinai and Fulda terrain. And, exclusion of data from one Inf attack did result in higher correlations between variables. This data point was excluded because the unit was not successful in its mission, suggesting that other factors contributing to the degree of mission accomplishment should be identified and included with simulation outcomes as measures of BCG performance.

Finally, there is the issue of no mutually agreed upon criterion measure of BCG performance. Research should be directed at identifying such a measure, so that the relationship of measures such as performance ratings, intrastaff information flow, and battlefield outcomes to overall BCG performance can be assessed. When these relationships are determined, it would be possible to develop an additional set of performance measures, which could be used as part of a diagnostic/ feedback package for BCGs. This feedback could support the training needs of BCGs in computer-driven command and control training exercises by providing indications of BCG strengths and weaknesses.

In conclusion, some degree of validity can be attributed to simulation outcomes as measures of BCG performance in CATTS exercises. The components of those formulas, battlefield losses, were sensitive to changes in mission and unit type, and probably differences in pre-battle combat ratios. Results are particularly promising due to high correlations between simulation outcomes and controller ratings of BCG performance. Further research on the development of indices of simulated battle performance should include control or systematic manipulations of the above variables along with controller influence and terrain factors. Indices should include additional components of mission accomplishment to supplement the simulation outcomes used in the current research.

REFERENCES

Barber, H.F., and Kaplan, I.T. Battalion command group performance in simulated combat. ARI Technical Paper 353, March 1979. (AD A070 089).

Kaplan, I.T. Information flow in battalion command groups. ARI Technical Report 499, June 1980. (AD A109 469)

Kaplan, I.T., and Barber, H.F. Training battalion command groups in simulated combat: Identification and measurement of critical performance. ARI Technical Paper 376, June 1979. (AD A075 414).

Stevens, S.S. Psychophysics. John Wiley and Sons, New York, 1973.

TRASANA. Validation of the combined arms tactical training simulation. TRASANA Technical Report 6-79, August 1979.

USACDC. Force developments: The measures of effectiveness. USACDC Phamphlet No. 71-1, January 1973.

APPENDIX A

Determination of Initial Strength

Initial levels of OPFOR were determined by summing the assets of all first echelon maneuver units and available artillery, regardless of mission and whether or not they actually became engaged in battle. Air assets and air defense weapons were not included, since they had no combat effectiveness weights in the CATTS model. However, both OPFOR and friendly aircraft inflicted casualties that were included in total losses.

Initial levels of blue strength included the maneuver strength of units under the direct control of BCG, plus adjacent unit strength when it was engaged in battle. In all scenarios, BCG units performed attack and defense missions in conjunction with adjacent, brigade controlled battalions. For example, Inf 1 and Inf 8 performed a Sinai static defense while adjacent battalions performed a covering force mission. In attack scenarios, BCG controlled forces were also assisted to differing degrees by adjacent battalions. In Cav attack scenarios, Cav units performed a screening mission for adjacent units which occasionally became engaged with the OPFOR. In all the above situations, adjacent units received and inflicted casualties on the OPFOR. Since the impact of adjacent units was potentially significant in the above scenarios, the strength levels of these adjacent units were added to BCG controlled unit strength, and casualties inflicted on and by the OPFOR were included in total loss figures. Artillery and air strength were not included in initial strength levels for friendly units because these assets were not under the direct control of BCGs.

The fact that a majority of a motorized rifle regiment was included in the initial strengths of OPFOR when friendly forces were attacking or defending, had the effect of reducing simulation outcome values when friendly forces were attacking. This caused the distributions of simulation outcomes for defense and attack missions to overlap. This manipulation allowed a comparison between controller ratings and simulation outcomes across mission types.

APPENDIX B

Calculations of Simulation Outcomes*

$$\text{Loss Exchange Ratio} = \frac{\text{Total OPFOR Losses}}{\text{Total Friendly Losses}}$$

$$\text{Relative Exchange Ratio} = \frac{\text{Percentage of OPFOR Lost}}{\text{Percentage of Friendly Forces Lost}}$$

Surviving Maneuver Force Ratio Differential = Percentage of Friendly Forces Surviving <u>minus</u> the Percentage of OPFOR Surviving

C^2ILL Ratio = 1/2 (Percentage of Friendly Forces Surviving) <u>plus</u> the Percentage of OPFOR Lost

$$\Delta\text{CR Combat Ratio} = \frac{\text{Initial Combat Ratio } \underline{\text{minus}} \text{ Ending Combat Ratio}}{\text{Initial Combat Ratio}}$$

* All losses are based on ΣEW X ET per exercise where EW = equipment weighting factor and ET = equipment type.

APPENDIX C

Battle Performance Indices

Covering Force/Defense

Unit	LER	RER	SMFRD	ΔCR	C^2ILL	Controller Performance Ratings in Z Scores
Mech 4	1.490	.498	-.213	-.371	.4994	-.2092
Mech 5	1.302	.451	-.264	-.464	.4764	-.5519
Mech 6	1.011	.341	-.282	-.495	.4320	-.8542
Mech 10	.980	.331	-.304	-.553	.4235	-1.3302
Cav 2	1.181	.623	-.198	-.414	.5644	.5096
Cav 3	1.121	.502	-.277	-.621	.5009	.4003
Cav 7	1.309	.645	-.182	-.375	.5744	.6986
Inf 1	.992	.639	-.168	-.314	.5647	.5558
Inf 8	.715	.622	-.191	-.383	.5616	.4918
Inf 9	1.070	.500	-.250	-.917	.4998	.1136

Attack

Unit	LER	RER	SMFRD	ΔCR	C^2ILL	Controller Performance Ratings in Z Scores
Mech 4	1.319	.608	-.108	-.148	.5298	.7128
Mech 5	.235	.113	-.281	-.408	.3777	-.1509
Mech 6	.623	.296	-.136	-.197	.4606	.2226
Mech 10	.743	.353	-.278	-.328	.4370	-.1065
Cav 2	.340	.310	-.225	-.405	.4298	.0589
Cav 3	.492	.465	-.175	-.261	.4888	.3896
Cav 7	.385	.363	-.107	-.128	.4771	.3395
Inf 1	.779	.857	-.068	-.131	.6704	-.8177
Inf 8	.452	.610	-.089	-.115	.5251	.6321
Inf 9	.710	.313	-.129	-.174	.4529	.0740